SOME ALGEBRAIC ANALOGIES IN MATRIC THEORY.

By Albert A. Bennett.

An obvious analogy exists between the theory of matrices and the theory of algebraic numbers. The analogy is in some respects superficial, but it is suggestive and extends further than is usually pointed out. A conspicuous cause of difference in the two theories is that while multiplication among algebraic numbers is always commutative, this is not the case among square matrices of a given order. As a result, a matric equation with scalar coefficients when satisfied by a given matrix is satisfied also by all transforms of this matrix through nonsingular matrices. The number of nonsingular distinct roots cannot usually be finite.

In the following discussion the matrices considered will be assumed without further mention to be square matrices and all of the same order. Such theorems concerning matrices as are found in Bôcher's "Introduction to Higher Algebra" will be assumed without discussion. The term "conjugate" as applied to a matrix will not be used in the current sense of the transposed matrix, obtained by turning the given matrix over about its main diagonal and thus interchanging rows and columns. On the contrary by "conjugate" will be meant the algebraic analogue of the term as used in the theory of algebraic numbers and given for matrices explicitly in detail by H. Taber.* The term "scalar" will be applied to a matrix having zeros except in the main diagonal and having the elements in the main diagonal equal. The "latent roots" of a matrix, or roots of the characteristic equation of a matrix will be called the *characteristic numbers* of the matrix.

Some Theorems Concerning Matrices Which Have Immediate Algebraic Analogues.

We shall list below a set of twenty-eight propositions concerning matrices, each of which may be translated at once into its counterpart in the theory of algebraic numbers. To do this it is merely necessary to substitute as follows:

For "identical matrix," I, substitute "unity" (1).

For "null matrix," 0, substitute "zero," 0.

For "scalar," substitute "rational number."

* H. Taber, On certain identities in the theory of matrices. Amer. Journ. Math., vol. 13 (1891), pp. 159–172.

For "number," substitute "integer."

For "matrix," substitute "algebraic number."

For "matrix with distinct non-vanishing characteristic numbers," substitute "Galoisian algebraic numbers."

For "characteristic function," substitute "defining function."

For "determinant," substitute "norm."

1. The identical matrix, I, and the null matrix, 0, are scalars.

2. Addition, subtraction, multiplication and division according to the usual rules of algebra may be performed among scalars.

3. The matric equation $ax = bI$ where a and b are numbers, and a is not zero, has a unique scalar as a solution, and each scalar is the root of such an equation.

4. If α is a non-scalar matrix there exists a polynomial

$$f(x) = x^n - s_1x^{n-1} + \cdots + (-1)^n s_n,$$

with scalar coefficients, of which α is a root.

5. There is a minimum degree (> 1) for such a function and there is but one function of this minimum degree.

6. Certain matrices are distinguished by many simple properties and are worthy of special study. For the present, only matrices with distinct nonvanishing characteristic numbers will be discussed, although some of the relations mentioned apply to all matrices.

7. The minimum degree n of the $f(x)$ for a matrix, α, of distinct non-vanishing characteristic numbers is called the *order* of α, and $f(x)$, *its characteristic function.*

8. The characteristic function, $f(x)$, of a matrix α of distinct non-vanishing characteristic numbers has a set of n distinct roots, $\alpha, \alpha_1, \alpha_2, \cdots, \alpha_{n-1}$, where $\alpha_1, \alpha_2, \cdots, \alpha_{n-1}$ are called the *conjugates* of α, and these satisfy the following conditions:

(*i*) Each conjugate, α_i, may be expressed as a polynomial in α with scalar coefficients.

(*ii*) Each conjugate, α_i, is a matrix of the same order, n, and with the same characteristic function, $f(x)$, as α.

(*iii*) The elementary symmetric functions of the set $(\alpha, \alpha_1, \cdots, \alpha_{n-1})$ are (except for sign) the n scalar coefficients $s_1, s_2, \cdots, s_n$ of the characteristic function,

$$f(x) = x^n - s_1x^{n-1} + \cdots + (-1)^n s_n.$$

9. The coefficient s_1 is called the *trace* of α, and the coefficient s_n, the *determinant* of α.

10. The function $f(x)$ may be viewed as the determinant of $(x - \alpha)$.

11. For α, a matrix with distinct nonvanishing characteristic numbers,

it is possible to select in many ways a *basis* of n matrices $\beta_1, \beta_2, \cdots, \beta_n$, linearly independent polynomials in α, with scalar coefficients, such that the totality of linear combinations with scalar coefficients, of the matrices of the basis, include all rational functions of α, where the indicated division has a meaning.

12. Two possible choices of a basis are

$$(1, \alpha, \alpha^2, \cdots, \alpha^{n-1}) \qquad \text{and} \qquad (\alpha, \alpha_1, \alpha_2, \cdots, \alpha_{n-1}).$$

13. In particular, for α, a matrix with distinct nonvanishing characteristic numbers, every rational function of α, where the indicated division results in a finite matrix and where the coefficients are scalars, is expressible as a polynomial in α of degrees less than n, with scalar coefficients.

14. The totality of such rational functions of α may be called the *domain* of α. Multiplication within the domain is commutative.

15. For any n matrices, $\gamma_1, \gamma_2, \cdots, \gamma_n$, of the domain of a matrix α of distinct nonvanishing characteristic numbers, the discriminant of $(\gamma_1, \gamma_2, \cdots, \gamma_n)$ is defined as the determinant of scalars,

$$\begin{vmatrix} S(\gamma_1\gamma_1) & S(\gamma_1\gamma_2) & \cdots & S(\gamma_1\gamma_n) \\ S(\gamma_2\gamma_1) & S(\gamma_2\gamma_2) & \cdots & S(\gamma_2\gamma_n) \\ \cdot & \cdot & \cdot & \cdot \\ S(\gamma_n\gamma_1) & S(\gamma_n\gamma_2) & \cdots & S(\gamma_n\gamma_n) \end{vmatrix},$$

where $S(\xi)$ is the *trace* of ξ. The discriminant is denoted by the symbol,

$$\Delta(\gamma_1, \gamma_2, \cdots, \gamma_n).$$

16. If $\gamma_i = \Sigma_j r_{ij}\beta_j$, where r_{ij} is scalar, then $\gamma_i\gamma_k = \Sigma_j r_{ij}\beta_j \Sigma_l r_{kl}\beta_l = \Sigma_{jl}(r_{ij}r_{kl})(\beta_j\beta_l)$. But $S(r\delta) = rS(\delta)$, where r is scalar, and $S(\delta_1 + \delta_2) = S(\delta_1) + S(\delta_2)$ for $\delta, \delta_1, \delta_2$, any matrices of the domain.

$$\therefore\ S(\gamma_i\gamma_k) = \Sigma_{jl}(r_{ij}r_{kl})S(\beta_j\beta_l).$$

By reference to the rule for multiplication of determinants, we have

$$\Delta(\gamma_1, \gamma_2, \cdots, \gamma_n) = [\text{Det } (r_{ij})]^2 \Delta(\beta_1, \beta_2, \cdots, \beta_n).$$

17. The discriminant of the basis $(\alpha, \alpha_1, \cdots, \alpha_{n-1})$ is not zero. It is expressible as

$$\begin{vmatrix} \alpha & \alpha_1 & \alpha_2 & \cdots & \alpha_{n-2} & \alpha_{n-1} \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_{n-1} & \alpha \\ \alpha_2 & \alpha_3 & \alpha_4 & \cdots & \alpha & \alpha_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \alpha_{n-1} & \alpha & \alpha_1 & \cdots & \alpha_{n-3} & \alpha_{n-2} \end{vmatrix}^2.$$

18. Hence the discriminant of every basis of the domain is different from zero.

19. It is possible to find a matrix α, of order n, no restriction as to the characteristic numbers being imposed, such that the equation $x^2 = \alpha$ is not satisfied by any proper matrix of order n.

20. In order to render certain general matric theorems as to the existence of a matric equation universally valid, it is sometimes necessary to introduce an improper root, which may be viewed as the limit of a finite matrix, as a convenient parameter approaches infinity.

21. The product of the $n - 1$ conjugates of α is a matrix of the domain of α, called the *adjoint* of α, $A(\alpha)$.

22. The determinant of the adjoint is

$$AA_1A_2\cdots A_{n-1} = A(\alpha)A(\alpha_1)\cdots A(\alpha_{n-1}) = (\alpha\alpha_1\cdots\alpha_{n-1})^{n-1},$$

which is the $(n - 1)$st power of the determinant of α.

23. The adjoint of the adjoint of α is in the same manner equal to α times the $(n - 2)$nd power of the determinant of α.

24. The sum of the $(n - 1)$ conjugates of α is a matrix of the domain of α called the adjoint-trace of α, $T(\alpha)$.

25. The trace of the adjoint-trace of α is $(n - 1)$ times the trace of α.

26. The adjoint-trace of the adjoint-trace of α is α plus $(n - 2)$ times the trace of α.

27. If γ is any matrix of the domain of α, the adjoint of $l - \gamma$, where l is scalar, is a polynomial in l of degree $n - 1$ with the coefficients in the domain.

28. If γ is any matrix of the domain of α, the determinant of $l - \gamma$, where l is scalar, is a polynomial in l of degree n with scalar coefficients.

A Discussion of Improper or Limit Matrices.

Any square matrix may be obtained as the limit of a matrix with distinct nonvanishing characteristic numbers, and theorems for a general matrix may sometimes be obtained by passage to a limit from this restricted but important case. It is needless to insist that care must be exercised. The well-known theorem that all matrices commutative with respect to multiplication with a given matrix of distinct nonvanishing characteristic numbers are rational integral functions of the given matrix has sometimes been stated for the general matrix. The theorem is, however, false, as is seen by reference to the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Some of the elements of a matrix β, which is obtained from a given matrix α of distinct nonvanishing characteristic numbers, may become infinite as two of the characteristic numbers of α approach equality, or one approaches zero. The limit may lead, therefore, not to a proper matrix but to an improper or limit matrix containing infinite elements.

An explicit mention of a similar limiting case is found in the classical memoir by Frobenius.* On pages 43 and 44 is found the following:

I. Every substitution, U (of determinant, $+1$), which transforms into itself a symmetric form, S, of nonvanishing determinant and for which the determinant of $E + U$ vanishes, may be expressed in the form

$$U = \lim\ (h = 0),\ (S + T_h)^{-1}(S - T_h),$$

where T_h is an alternating form whose coefficients are rational functions of h.

II. Every substitution, U, which transforms into itself an alternating form, T, of nonvanishing determinant, and for which the determinant of $E - U$ vanishes, may be expressed in the form

$$U = \lim\ (h = 0),\ (S_h + T)^{-1}(S_h - T),$$

where S_h is a symmetric form whose coefficients are rational functions of h.

Another occasion for the use of improper matrices is in the extraction of square roots of matrices. While for e, different from zero, $\left(\begin{smallmatrix} e^2 & 2 \\ 0 & e^2 \end{smallmatrix}\right)$ has the square root $\left(\begin{smallmatrix} e & 1/e \\ 0 & e \end{smallmatrix}\right)$, yet for $e = 0$, there is no proper matrix obtained as a square root but only an improper matrix, as a limit.

The relations between the trace, adjoint-trace, determinant, and adjoint may be so expressed as to be valid for all square matrices without restriction as to characteristic numbers. Thus there are certain relations which in terms of the conjugates of a matrix become obvious but which are capable of proof without reference to conjugates. Many of the arguments which have resulted in the successive historical extensions of the number system and in the introduction as valid numbers of negatives, fractions, irrationals, imaginaries, may be urged for the acceptance of limit matrices, at least when these are required to render general the notion of conjugates.

The matrix, α, taken as

$$\begin{pmatrix} e & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

for $e \neq 1, \neq 2$, has two proper conjugates in the sense used above, which may be taken as

$$\begin{bmatrix} 2 & -\dfrac{2-e}{1-e} & 0 \\ 0 & e & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & \dfrac{1}{1-e} & 0 \\ 0 & 2 & 0 \\ 0 & 0 & e \end{bmatrix},$$

* Frobenius, Über lineare Substitutionen und bilineare Formen. Jour. f. d. reine und ang. Math., vol. 84 (1878), pp. 1–63.

which become improper as e approaches unity. For $e = 1$, there is no set of proper conjugates. It will not be sufficient to denote both lim $(e = 1)$, $-(2 - e)/(1 - e)$, and lim $(e = 1)$, $1/(1 - e)$ by the mere sign ∞. The algebraic relations between these quantities must be retained also in the limit. Despite these difficulties, symbols α_1 and α_2 may be used for these limit matrices and the correct relations may be found by their means among the quantities: trace, adjoint-trace, determinant and adjoint. It is merely necessary to regard α_1 and α_2 as not themselves in the domain of α, although commutative with α in multiplication and giving rise to the same characteristic functions. This is analogous to going from the Galois domains to non-Galois domains.